

J. Gjønnes for his interest in the present work. The labradorite specimens came from the collection of the Mineralogical–Geological Museum, Oslo.

APPENDIX

$$\begin{aligned}
 S_{11} &= b^2 c^2 \sin^2 \alpha, \\
 S_{22} &= c^2 a^2 \sin^2 \beta, \\
 S_{33} &= a^2 b^2 \sin^2 \gamma, \\
 S_{12} &= abc^2 (\cos \alpha \cos \beta - \cos \gamma), \\
 S_{23} &= a^2 bc (\cos \beta \cos \gamma - \cos \alpha), \\
 S_{13} &= ab^2 c (\cos \gamma \cos \alpha - \cos \beta), \\
 V^2 &= a^2 b^2 c^2 (1 - \cos^2 \alpha - \cos^2 \beta \\
 &\quad - \cos^2 \gamma + 2 \cos \alpha \cos \beta \cos \gamma), \\
 M^2 &= \frac{1}{d_{pqr}^2} = \frac{1}{V^2} (S_{11} p^2 + S_{22} q^2 \\
 &\quad + S_{33} r^2 + 2S_{12} pq + 2S_{23} qr + 2S_{31} pr).
 \end{aligned}$$

References

- BAMBAUER, H. U., EBERHARD, E. & VISWANATHAN, K. (1967). *Schweiz. Miner. Petrogr. Mitt.* **47**, 351–364.
- BOLLMANN, W. (1970). *Crystal Defects and Crystalline Interfaces*. Berlin: Springer.
- BOWN, M. G. & GAY, P. (1958). *Z. Kristallogr.* **111**, 1–14.
- CLIFF, G., CHAMPNESS, P. E., NISSEN, H.-U. & LORIMER, G. W. (1976). *Electron Microscopy in Mineralogy*, pp. 258–265. Berlin: Springer.
- FAIVRE, P. (1975). *J. Appl. Cryst.* **8**, 356–360.
- HASHIMOTO, H., KUMAO, A., ENDOH, H., NISSEN, H.-U., ONO, A. & WATANABE, E. (1976). *Developments in Electron Microscopy and Analysis*, pp. 245–250. New York: Academic Press.
- HASHIMOTO, H., NISSEN, H.-U., ONO, A., KUMAO, A., ENDOH, H. & WOENSDREGT, C. F. (1976). *Electron Microscopy in Mineralogy*, pp. 332–344. Berlin: Springer.
- HIRSCH, P. B., HOWIE, A., NICHOLSON, R. B., PASHLEY, D. W. & WHELAN, M. J. (1965). *Electron Microscopy of Thin Crystals*. London: Butterworths.
- HØIER, R. (1969). *Acta Cryst.* **A25**, 516–518.
- JOHARI, O. & THOMAS, G. (1969). *The Stereographic Projection and its Applications, Techniques of Metals Research*. Vol. Ila. New York: Interscience.
- MCLAREN, A. C. & MARSHALL, D. B. (1974). *Contr. Miner. Petrol.* **44**, 237–249.
- NISSEN, H.-U. & BOLLMANN, W. (1968). *Proc. 4th Eur. Reg. Conf. Electron Microscopy, Rome*, pp. 321–322.
- NISSEN, H.-U., CHAMPNESS, P. E., CLIFF, G. & LORIMER, G. W. (1973). *Nature, Lond.* **245**, 135–137.
- OLSEN, A. (1976a). *J. Appl. Cryst.* **9**, 9–13.
- OLSEN, A. (1976b). *Proc. 6th Eur. Conf. Electron Microscopy, Jerusalem*, Vol. 1, pp. 417–418.
- OLSEN, A. & LILLEBÖ, V. (1976). *Developments in Electron Microscopy and Analysis*, pp. 489–492. New York: Academic Press.
- PIROUZ, P. & BOSWARVA, I. M. (1974). *Phys. Stat. Sol. (a)*, **26**, 407–415.
- PUMPHREY, P. H. & BOWKETT, K. M. (1970). *Phys. Stat. Sol. (a)*, **2**, 339–346.
- SMITH, J. V. (1974a). *Feldspar Minerals*, Vol. I. New York: Springer.
- SMITH, J. V. (1974b). *Feldspar Minerals*, Vol. II. New York: Springer.
- STEWART, D. B. & WRIGHT, T. L. (1974). *Bull. Soc. Fr. Minér. Crist.* **97**, 356–377.

Acta Cryst. (1977). **A33**, 712–716

Solving Structures with Quartets: The Least-Squares Analysis of Quartet Invariants in Space Group $P1$

BY CHRISTOPHER J. GILMORE

Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland

(Received 26 March 1977; accepted 29 April 1977)

A least-squares technique for extracting individual phase angles from a set of quartet invariants is described. For symmorphic space groups, this procedure offers the advantages of stability and, in non-centrosymmetric cases, a systematic way of defining the enantiomorph, in contrast with traditional direct methods employing triplets. The application to a phthalic anhydride derivative $C_{26}H_{16}O_5$ in space group $P1$ is described. The method is readily extended to other space groups.

1. Introduction

Crystal structures in space group $P1$ are traditionally the most difficult to solve by direct methods. The symmorphic nature of the space group tends to make the process ill-conditioned, and the lack of equivalent reflexions gives rise to a paucity of sign relations of low

associated variance. This forces the need for a relatively large starting set which, in turn, implies a large number of possible solutions from which the correct phase set may be difficult to extract. Enantiomorph definition is often a haphazard affair, since it is difficult to predict accurately those invariants with magnitudes sufficiently far from 0 or π .

Recent formulae derived by Hauptman (1975, 1976) and Giacovazzo (1976a) for estimating quartet invariants in space group $P1$ offer certain advantages over the conventional distribution for triple-phase invariants. In $P1$, quartets are structure invariants of the form:

$$\Phi = \varphi_h + \varphi_k + \varphi_l + \varphi_m \quad (1.1)$$

where:

$$\mathbf{h} + \mathbf{k} + \mathbf{l} + \mathbf{m} = 0 \quad (1.2)$$

and $\varphi_h, \varphi_k, \varphi_l$, and φ_m are the phase angles associated with the normalized structure factors E_h, E_k, E_l, E_m . Each quartet has an associated variable

$$B = \frac{3\sigma_3^2 - \sigma_2\sigma_4}{\sigma_2^3} |E_h E_k E_l E_m| \quad (1.3)$$

where

$$\sigma_j = \sum_{i=1}^N z_i^j. \quad (1.4)$$

There are N atoms in the unit cell with atomic number z_i . The $P_{1|7}$ formula (Hauptman, 1976) estimates Φ as a function of the four principal E magnitudes $|E_h|, |E_k|, |E_l|$ and $|E_m|$ (R_1, R_2, R_3, R_4 respectively) plus the three unique cross-terms $|E_{h+k}|, |E_{k+l}|, |E_{l+h}|$ (R_{12}, R_{23}, R_{31} respectively). The formula is couched in terms of the conditional joint probability distribution

$$\begin{aligned} P(\Phi | R_1, R_2, R_3, R_4, R_{12}, R_{23}, R_{31}) \\ \simeq \frac{1}{M} \exp(-2B \cos \Phi) I_0 \left(\frac{2\sigma_3}{\sigma_2^{3/2}} R_{12} Y_{12} \right) \\ \times I_0 \left(\frac{2\sigma_3}{\sigma_2^{3/2}} R_{23} Y_{23} \right) I_0 \left(\frac{2\sigma_3}{\sigma_2^{3/2}} R_{31} Y_{31} \right) \end{aligned} \quad (1.5)$$

where:

$$Y_{12} = (R_1^2 R_2^2 + R_3^2 R_4^2 + 2R_1 R_2 R_3 R_4 \cos \Phi)^{1/2} \quad (1.6)$$

$$Y_{23} = (R_2^2 R_3^2 + R_1^2 R_4^2 + 2R_1 R_2 R_3 R_4 \cos \Phi)^{1/2} \quad (1.7)$$

$$Y_{31} = (R_3^2 R_1^2 + R_2^2 R_4^2 + 2R_1 R_2 R_3 R_4 \cos \Phi)^{1/2}. \quad (1.8)$$

M is a normalizing constant. The distribution is defined in terms of $\cos \Phi$ and is therefore unique only in the range $0 - \pi$. Put another way, only the magnitude of the quartet invariant is available and not the sign. Fig. 1 shows the distribution (1.5) when both principal and cross-terms are large; in this case Φ has a modal value of zero. For large principal terms and small cross-terms Φ has a most probable value of π (Fig. 2). The availability of quartets with probable values of π (negative quartets) alleviates the symmorphic nature of the phase problem in space group $P1$. A final case of large principal terms and medium-sized cross-terms is shown in Fig. 3. Here Φ has a magnitude of approximately $\pi/2$ and this quartet can be used as a possible enantiomorph definer. All three distributions are derived from the crystal structure analysis described in

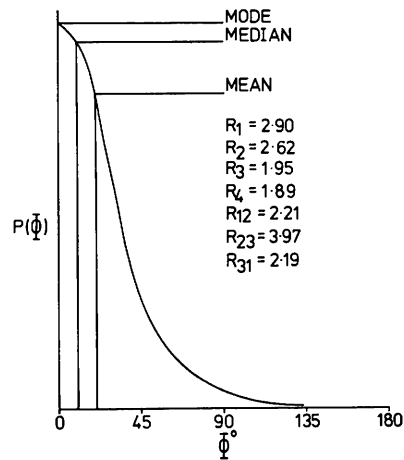


Fig. 1. The distribution (1.5) for the E magnitudes shown. The standard deviation of the mean is 17° .

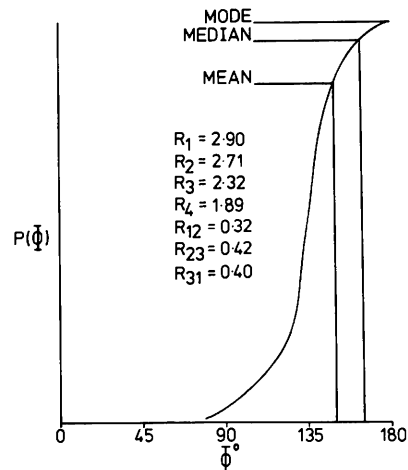


Fig. 2. The distribution (1.5) for the E magnitudes shown. The standard deviation of the mean is 21° .

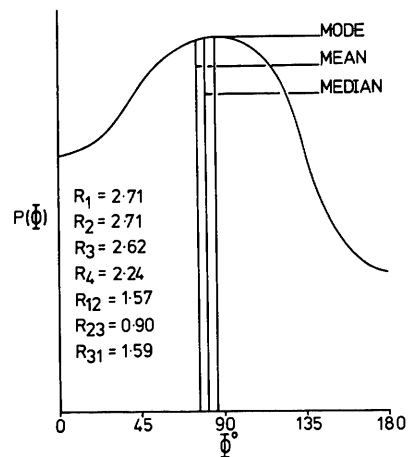


Fig. 3. The distribution (1.5) for the E magnitudes shown. The standard deviation of the mean is 39° .

§4. The examples quoted are unimodal, and experience would indicate that this is always the case.

For a given set of conditions there is usually an order of magnitude more quartets than triplets, so that there is no longer a paucity of phase relations.

2. Statistical parameters derivable from the conditional joint probability distribution

As with other conditional joint probability distributions several measures of central tendency are available. (a) The mode $|\Phi|_m$, which is the most probable Φ magnitude. (b) The mean, $\langle|\Phi|\rangle$, derived from

$$\langle|\Phi|\rangle = \int_0^\pi \Phi P(\Phi) d\Phi \quad (2.1)$$

where $P(\Phi)$ is an abbreviation of the left hand of (1.5).

(c) The median, $|\Phi|_{md}$, which for conditions of small skewness of the distribution is derived as

$$|\Phi|_{md} = \langle|\Phi|\rangle - 1/3(\langle|\Phi|\rangle - |\Phi|_m). \quad (2.2)$$

These parameters are independent of M . However, it is necessary to have a measure of the variance of the mean (V) and its associated standard deviation (σ) for which M is needed. Hauptman (1975) has derived an analytic expression for M , but it is computationally simpler to derive it numerically such that

$$\int_0^\pi P(\Phi) d\Phi = 1. \quad (2.3)$$

V is then:

$$V = \int_0^\pi (\Phi - \langle|\Phi|\rangle)^2 P(\Phi) d\Phi \quad (2.4)$$

and

$$\sigma = V^{1/2}. \quad (2.5)$$

All these parameters are readily derived by numerical integration on any modern computer and require only small amounts of central-processor time.

3. The least-squares analysis of quartet invariants

The problem now arises as to how best to utilize this information in a phase-determining procedure. Since the problem of deriving the individual phase angles from the quartet invariants is greatly overdetermined, and since only Φ magnitudes are available, a least-squares procedure similar to that described by Hauptman (1972) suggests itself.

If a series of cosine invariants of the form:

$$\cos(\varphi_h + \varphi_{k1} + \varphi_{l1} + \varphi_{m1}) \simeq c_1 \quad (3.1)$$

$$\cos(\varphi_h + \varphi_{k2} + \varphi_{l2} + \varphi_{m2}) \simeq c_2 \quad (3.2)$$

$$\cos(\varphi_h + \varphi_{kn} + \varphi_{ln} + \varphi_{mn}) \simeq c_n \quad (3.3)$$

is available in which the coefficients c_1, c_2, \dots, c_n and the phase angles $\varphi_{k1}, \varphi_{k2}, \dots, \varphi_{l1}, \varphi_{l2}, \dots, \varphi_{m1}, \varphi_{m2}, \dots$ are all known, at least approximately, then the common

phase angle φ_h may be determined by minimizing the function

$$\psi = \sum_j \omega_j [\cos(\varphi_h + \varphi_{kj} + \varphi_{lj} + \varphi_{mj}) - c_j]^2 / \sum_j \omega_j \quad (3.4)$$

where ω_j is a weight associated with each quartet. In practice it is set to the inverse of the variance derived from (2.4). (3.4) offers an alternative to the standard triple-phase determining formulae used in most direct-methods procedures although the right hand of (3.4) need not consist of quartets exclusively but any invariant or seminvariant for which a cosine estimate is available. Thus triple-phase invariants with large associated A values can also be included with cosine values of 1.0 or with cosines estimated by the TPROD or MDKS formulae (Hauptman, 1972).

As with all phase-determining procedures, a measure of the reliability with which φ_h is determined is required. A simple measure is the depth of the minimum of ψ :

$$d = \psi_{\max} - \psi_{\min}. \quad (3.5)$$

The maximum value of d is 4.0. However, this takes no account of the number of contributors, n , to the minimum; by trial this seems best included with d to give a modified figure of merit of the form $dn^{1/2}$. A further complication arises in early stages of phase determination when only single invariants are being used to derive new phase angles. The best-determined phases should come from those invariants with the lowest associated standard deviations. We therefore append to the modified figure of merit, the weight, ω_{\max} , associated with the most reliable invariant used to generate the phase angle φ_h , to obtain the function

$$R_h = \omega_{\max} dn^{1/2}. \quad (3.6)$$

High values of R_h are associated with those phase angles which have been reliably determined. Typically, R_h is greater than 0.01 for such phases.

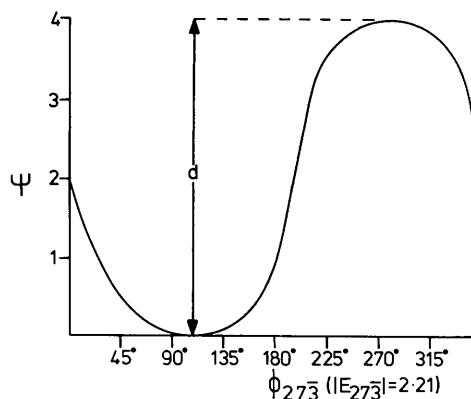


Fig. 4. The variation of ψ (3.4) as a function of φ_{273} . There are eight contributors, $d = 4.0$, $\omega_{\max} = 0.0027$, $R_h = 0.03$, $|E_{273}| = 2.21$.

Fig. 4 shows the behaviour of ψ for a well-determined phase having an associated value of $R_h = 0.03$. In contrast, Fig. 5 shows a poorly defined minimum in ψ with an associated value of $R_h = 0.004$.

Multiple minima in ψ occurring at values of ϕ_h sufficiently removed from 0 or π may be used for enantiomorph definition.

4. Application to the solution of the crystal structure of 6-acetoxy-7,9-diphenylcyclohepta[*c,d*]isobenzofuran-2,8-dione

$C_{26}H_{16}O_5$ is a phthalic anhydride derivative which crystallizes in space group *P1* with $Z=1$ (Freer, Gilmore, Mant & McCormick, 1977). Numerous attempts to solve this structure with vector search techniques and conventional direct methods were unsuccessful.

Normalized structure factors were calculated incorporating the molecular scattering factor for benzene (Main, 1976), and 8668 quartets with $B \geq 0.9$ were estimated from (1.5), for 231 *E* magnitudes > 1.57 . For each quartet the magnitudes of the mean, mode and median were computed with the associated variance and standard deviation. This procedure used *ca* 5 min of c.p.u. time on an IBM 370/168 computer. The 4000 quartets with the lowest variance were input into the least-squares program.* Three reflexions, $1\bar{1}1$, $1\bar{2}1$ and $1\bar{9}6$, were used to fix the origin. For a satisfactory phasing path two further phases were required, and the 048 and the $2\bar{1}\bar{1}0$ reflexions were also incorporated into the starting set.

In order to use a quartet relation to determine a new phase angle, three phase angles must be known. If these known phases are in error, a rapid propagation of errors is possible and is likely to be worse than the corresponding procedure that employs triplets, where only two known phases are required to generate a

* Programs *QGEN* (generation of quartets) and *QLS* (least-squares analysis) available on request.

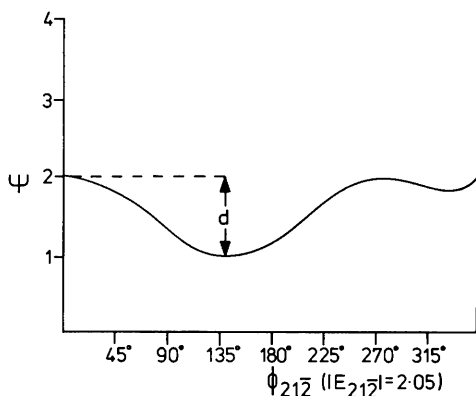


Fig. 5. The variation of ψ (3.4) as a function of ϕ_{212} . There are five contributors, $d = 1.3$, $\omega_{\max} = 0.0014$, $R_h = 0.004$, $|E_{212}| = 2.05$.

third. Whereas for triplets, phase-angle permutations of $\pm \pi/4$, $\pm 3\pi/4$ (or $0, \pi, \pm \pi/2$) may be sufficient (*viz* increments of $\pi/2$) for starting-set reflexions having general phases, this was found to be too coarse for use with quartets in space group *P1* and increments of $\pi/4$ from 0 to $7\pi/4$ were used. Suitable multiple minima in ψ for use in enantiomorph definition were not readily available in the early stages of least-squares analysis. Therefore, to define the enantiomorph and reduce the number of solutions to be investigated the value of the phase angle of reflexion 048 was constrained to the initial values $\pi/4, \pi/2, 3\pi/4$. The phase of the $2\bar{1}\bar{1}0$ reflexion was given eight possible initial values from zero to $7\pi/4$ in increments of $\pi/4$.

The necessity of defining three phases to give a fourth also gives difficulties in getting started in the phase-determination procedure. If quartets alone are used, a large starting set is often required. Relations between the starting-set phases quickly appear and hence reduce the number of possible solutions, but in a relatively time-consuming least-squares environment it is better to limit the starting set to the smallest compatible with a well-behaved phase-determining path. To this end, 109 unique triple-phase invariants having $A \geq 5.0$ were included in the analysis with an assumed cosine of +1 and a variance derived from the Cochran (1955) distribution by (2.4).

As with the least squares of triple-phase invariants, practical problems arise concerning the number of phases to be determined in a given cycle and whether these are to be held fixed once calculated. In the final successful procedure, 20 cycles of least-squares were performed initially for each starting set accepting only one new phase ϕ_h having the largest value of R_h in that cycle. The minimum in ψ was found numerically by varying ϕ_h in a coarse 30° grid from 0 to 330° and then stepping in 5° increments around the minima thus found. Two or more minima in ψ were assumed to exist if the corresponding values of ψ differed by less than 20%. A further 15 cycles were then carried out accepting the best 3, 5, 7, ... phases in each subsequent cycle. Once determined, a new phase was not refined unless the new value of R_h was greater than the previous value. A maximum shift of 70° of any phase in any cycle was imposed. At the end of each solution all the phases were refined together for one cycle. In contrast with a simple application of the tangent formula, such a procedure seemed to be stable for the compound studied. The cosine of the mode was used in (3.4). Similar studies with the mean or the median were unsuccessful. On an IBM 370/168 computer this method utilized *ca* 30 s of c.p.u. time per solution. *E* maps were produced directly from these phase sets without further refinement with the tangent formula.

5. Selecting the correct solution

The above method produced 24 possible solutions from which we wish to select the best set. The tradi-

tional figures of merit R_{Karle} (Karle & Karle, 1966) and ABSFOM (Germain, Main & Woolfson, 1971) are unreliable in symmorphic space groups, being highly dependent on the ill-behaved tangent formula. Two alternative figures of merit were employed in the least-squares procedure: (i) $\sum_{\mathbf{h}} R_{\mathbf{h}}$ which is a measure of consistency of the phase relations and should be a maximum for the correct solution; (ii) NQUEST (De Titta, Edmonds, Langs & Hauptman, 1975), the negative quartet figure of merit defined as:

$$\text{NQUEST} = \sum_{\mathbf{hklm}} \omega_{\mathbf{hklm}} \cos(\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} + \varphi_{\mathbf{m}}) / \sum_{\mathbf{hklm}} \omega_{\mathbf{hklm}} \quad (5.1)$$

where the summations are taken over all those quartets predicted to have a modal value of π . The use of this formula for the centrosymmetric case has already been discussed (Gilmore, Hardy, MacNicol & Wilson, 1977) and the points raised there are equally valid in the non-centrosymmetric situation.

With the existence of a reliable conditional joint probability distribution, it is possible to remove the arbitrary constraints imposed initially on the summations in (5.1) by summing over all quartets having estimated cosines < 0 , setting $\omega_{\mathbf{hklm}} = 1/\sigma_{\mathbf{hklm}}^2$ and including no B limit. A theoretical value of NQUEST (N_{calc}) is then derived by substituting the quartet cosine estimates from (1.5) into (5.1). An observed value of NQUEST (N_{obs}) can also be derived by substituting the calculated phase angles into (5.1) and summing over the same quartet invariants. A joint figure of merit is given by:

$$N_c = \frac{1}{2}(N_{\text{obs}} + N_{\text{calc}}). \quad (5.2)$$

The optimum value of N_c is -1.0 where $N_{\text{obs}} = N_{\text{calc}} = -1$. Complete disagreement between the phasing path and quartet theory gives $N_c = 0$.

These two figures of merit were computed for all 24 solutions. The solution having $N_c = -0.46$ and $\sum R_{\mathbf{h}} = 11.3$ yielded an E map in which the positions of 21 of the 31 non-hydrogen atoms were revealed. The structure was completed by weighted Fourier techniques (Sim, 1960) and refined to $R = 0.107$.

The two figures of merit and the unmodified NQUEST showed a disappointing lack of discriminating power; indeed all the solutions were in the range of -0.56 to -0.42 for N and 11.3 to 11.9 for $\sum R_{\mathbf{h}}$. In view of the previous experiences of De Titta *et al.* (1975) this may seem surprising. However, as mentioned for the centrosymmetric case (Gilmore *et al.*, 1977) previous successful applications of this figure of merit have not used the negative quartets for *a priori* phasing, but have utilized them passively at the end of the analysis as a figure of merit that is independent of the phase-determining procedure. This independence is lost when the quartets are used in an active way, and the reliability of the calculation reduced. However, it is anticipated that, when quartets are not used for *a priori*

phasing, this modified formulation of NQUEST will be both easier to calculate and more reliable.

6. Application to space groups of higher symmetry

Although the P_{117} formula (1.5) was derived specifically for space group $P1$, it should be valid for conditions of higher symmetry (Giacovazzo, 1976*b*). Probabilistic formulae also exist for quartet invariants in the centrosymmetric case (Giacovazzo, 1976*c*; Hauptman, 1976). The least-squares technique can be applied to these situations in the same way. An additional acceptance criterion is required for reflexions having restricted phases, where it is necessary to ensure that the minimum in (3.4) is within a suitable distance of the relevant restriction. Those phases for which this is not so are poorly determined and should be rejected.

In most cases of higher symmetry, the problem of symmorphism is not encountered, yet the tangent formula can still sometimes be unstable (Lessinger, 1976). The least-squares technique then offers an alternative which is more likely to be stable. In larger structures there is often a paucity of reliable phase relations when triplets alone are used. The incorporation of quartets increases this number, and the least-squares analysis provides a flexible method of utilizing them as well as permitting the incorporation of any other invariant or seminvariant. We are currently exploring the use of this procedure in these cases in our laboratory.

The author thanks Dr H. Hauptman for useful discussions, Dr A. F. Cameron for the intensity data for the compound studied, and Mr A. Freer who assisted in the computational procedures.

References

- COCHRAN, W. (1955). *Acta Cryst.* **8**, 473–478.
 DE TITTA, G. T., EDMONDS, J. W., LANGS, D. A. & HAUPTMAN, H. (1975). *Acta Cryst.* **A31**, 472–479.
 FREER, A. A., GILMORE, C. J., MANT, D. M. & MCCORMICK, J. (1977). *Chem. Commun.* 296–297.
 GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* **A27**, 368–376.
 GIACOVAZZO, C. (1976*a*). *Acta Cryst.* **A32**, 91–99.
 GIACOVAZZO, C. (1976*b*). *Acta Cryst.* **A32**, 958–966.
 GIACOVAZZO, C. (1976*c*). *Acta Cryst.* **A32**, 74–82.
 GILMORE, C. J., HARDY, A. D. U., MACNICOL, D. D. & WILSON, D. R. (1977). *J. Chem. Soc. Perkin II*. In the press.
 HAUPTMAN, H. (1972). *Crystal Structure Determination: The Role of the Cosine Seminvariants*. New York and London: Plenum.
 HAUPTMAN, H. (1975). *Acta Cryst.* **A31**, 680–687.
 HAUPTMAN, H. (1976). *Acta Cryst.* **A32**, 877–882.
 KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.
 LESSINGER, L. (1976). *Acta Cryst.* **A32**, 538–550.
 MAIN, P. (1976). *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 97–114. Copenhagen: Munksgaard.
 SIM, G. A. (1960). *Acta Cryst.* **13**, 511–512.